ARCHIVES OF ACOUSTICS
15, 3–4, 325–348 (1990)

PERCEPTION OF POLISH ACCENT IN A RE-SYNTHESIZED SPEECH SIGNAL

W. JASSEM

Acoustic Phonetics Department, Institute of Fundamental Technological Research
Polish Academy of Sciences
(61-704 Poznań, Noskowskiego 10)

W. VAN DOMMELEN

Institut für Phonetik und digitale Sprachverarbeitung
Christian-Albrechts-Universität •Kiel
(Kiel, Bundesrepublik Deutschland)

The utterance /vidzîtɕeteras/ was spoken, in three versions, with the accent (I) on the first, (II) on the second and (IV) on the fourth syllable. Using a package of computer programs for speech processing which executes analysis and synthesis by linear prediction, the fundamental frequency parameter, or F_0, was mutually exchaned among the three versions, in two listening experiments. In further experiments, the time variations of F_0 were averaged in several ways, and replaced either by random noise or by a constant F_0 value. The listening experiments were carried out with three groups of subjects, viz. (1) German University students, (2) Polish University students, and (3) Polish research workers and technicall staff. The listeners in the first group, who knew no Polish, but specialized in phonetics, placed accent almost exactly like those of the Polish subjects whose reactions were systematic. The temporal variations of the F_0 parameter had a decisive effect on the perception of accent, whilst the effect of the other parameters was extremely limited when F_0 was replaced by random noise. Substitution of constant F_0 made the localization of accent almost totally impossible.

1. Introduction

The acoustical nature of speech accent belongs to the most controversial problems of contemporary experimental phonetics. An academic textbook of phonetics, still regarded as standard (LADEGOGED [9]) contains the following

definition of accent: "A stressed (i.e. accented, W. J.) syllable is pronounced with greater amount of energy than an unstressed syllable." (p. 222). In a terminological glossary in the same work, the description of accent (or stress) also relies on "extra respiratory energy during a syllable" (p. 281). Ladefoged adduces no experimental support, not even by way of a convincing illustration, for his concept of accent (or, using his term, "stress"), which – by the way – is widely accepted in linguistic circles.

Over the last thirty years or so, three general descriptions of contemporary Polish pronunciation have appeared: WIERZCHOWSKA ([16, 17, 18]). In all these works, the Polish accent (stress) is defined as "dynamic" or "expiratory", implying and increase of "force" (or vocal effort) within that part of an utterance which bears accent (WIERZCHOWSKA [16], p. 157; [17], p. 17; [18], p. 133–134). Translating this into acoustic terms, it should be expected that an accented syllable, other things being equal, would be signalled by relatively high intensity level compared with an unaccented (unstressed) syllable.

About thirty year ago, one of the present authors (JASSEM 1959, 1962), using measurements of three time-variable acoustic-phonetic parameters – fundamental frequency, intensity level envelope and segmental duration – attempted to show that the location of accent in Polish is essentially determined by the temporal variations of F_0 (not necessarily its maximum). Subsequently, JASSEM, MORTON and STEFFEN-BATÓG [8] presented the results of their experiments, performed with speech-like stimuli, which appeared to demonstrate that the F_0 parameter was decisive for the perception of accent (stress) by Polish listeners. Yet, 10 years later, DOBROGOWSKA [3] stated, on the basis of F_0 measurements, that this parameter was not unambiguously related to speech accent. The problem of the **acoustic** nature of accent (stress) thus remained open. The present investigation attacks this issue once again as it is currently especially acute in view of the needs of **speech synthesis** and computerized **recognition of the speech signal**.

A side-product of the study (not, apparently, devoid of some general interest), is the finding that in perceptual experiments in which acoustic-phonetic features are being examined, it is necessary to ensure homogeneity of subects' responses.

2. Potential and real accent

In the phonetic literature, the traditional terms "word stress" and "sentence stress" are still very much in use. Relating as they do to a specified syllable within a word or a sentence, they are not appropriate, primarily because these two linguistic units are not phonetic entities. They belong to morphology and syntax, and therefore to a different plane of analysis and description. JASSEM and GIBBSON [7] proposed the introduction into phonetic and phonological analysis of the English terms "stress" and "accent" in a sense of essentially new (though related to Bolinger's work:

1955a, b; 1958) definitions. The notions are independent of the word and the sentence, but are associated with them – more strictly, with the word and the intonation phrase respectively. Stress was defined as the **potential location** of accent, whilst the latter as the **real** fact of a syllable standing out in some specfic way. In the special case of English, certain durational relations and certain F_0 features determine accent.

In the present study we shall use the terms "stress" and "accent" as defined in JASSEM and GIBBON [7].

3. Experimental material

The utterance /vidzi͡tceteras/ was spoken by WJ (one of the present authors) in a sound-proof studio, and tape-recorded with high-quality instruments, in three versions, varying the location of accent, which fell, in version I on the first, in version II on the second and in version IV on the fourth syllable. These phrases may correspond either to the orthographic form **Widzi cię teraz** or **Widzicie teraz**. In either case, there are two stresses: on /vi/ and /te/ or on /dzi/ and /te/.

We used a package of programmes for the analysis and synthesis of speech developed in the Institute of Phonetics and Digital Speech Processing of the University of Kiel, which are based on linear predition and have been described in RATHJEN [13], Schäfer-Vincent 1972 and BARRY et al. [1]). The three original phrases were analyzed and processed in various ways decribed below, to obtain the output signals constituting our experimental material. This material was used in six experimental designs, each including a number of stimuli. Within each design, each stimulus was preceded by a brief 200 Hz beep of 100 ms duration and followed by a 4 s silent interval, this interval being later used by the listeners to make a decision as to the location of accent in the stimulus just heard. For the listeners' convenience, additional brief sinusoidal audio signals were recorded before every tenth stimulus.

Experiment 1

Three different stimuli were prepared: I_I – a re-synthesis of version I with the values of all analysis parameters unaltered, I_{II} – version I re-synthesized, with all parameters unaltered except F_0 which – after minor linear adjustments – was transplanted from version II, and I_{IV} – version I re-synthesized with all parameters unaltered except F_0 which – again after minor adjustments – was transplanted from version IV. Each of the three new stimuli: I_I, I_{II} and I_{IV} was copied 10 times to a total of 30 stimuli. Figure 1 shows the results of the LPC analysis of version I (with the first syllable accented). The phonetic parameters are: F_0, the frequency and bandwidth of F_1, F_2, F_3, F_4, and F_5, (the local spectrum level maxima) and time envelope of the intensity level (the LPC residue).

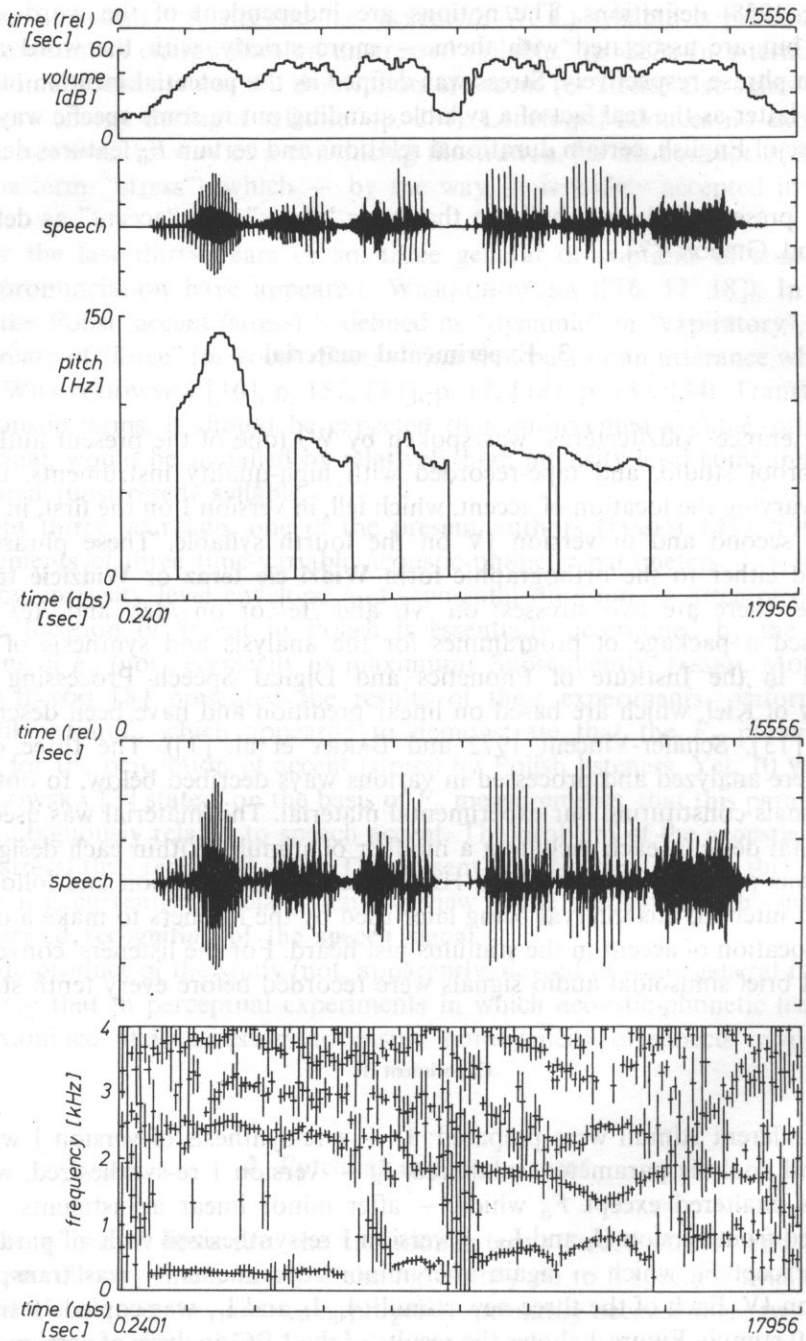


FIG. 1. The acoustic-phonetic parameters extracted by LPC from the utterance **Widzi cię teraz**

Experiment 2

In all the three original versions, the F_0 parameter was replaced by white noise. Each new stimulus was copied 10 times to a total of 30. The new stimuli were denoted as I_S, II_S, IV_S.

Experiment 3

The stimuli were here analogous to those in Experiment 1 with the addition of F_0 from utterance I and II transplanted into IV and utterances II and IV unaltered. As a result of these manipulations 9 different stimuli were obtained: I_I, I_{II}, I_{IV}; II_I, II_{II}, II_{IV}; IV_I,IV_{II} and IV_{IV}. Each of these was copied 10 times making up 90 stimuli in this experiment.

Experiment 4

The original versions were set in three pairs: (1) I_I and I_{II}, (2) I_I and I_{IV}, and I_{II} and I_{IV}. Within each pair, additional three intermediate versions were synthesized. These represented gradual transitions from the first to the second member of the pair with respect to F_0. Three sets of stimuli were thus produced, each including five items: the extreme ones (e.g. I_I and I_{IV}) and three intermediate ones. Each element of each set was denoted as in the following example: I_I, $(I_I{:}\, I_{IV})_2$, $(I_I{:}\, I_{IV})_3$, $(I_I{:}\, I_{IV})_4$, I_{IV}. As before, the Roman number on the line indicated which syllable was accented in the original, the subscript Roman number refers to the transplanted syllable, and the subscript Arabic number indicates the successive graded transitions from the initial to the final item. All the remaining parameters of the original remained unaltered, whilst the intermediate F_0 values were obtained, for each successive time interval of LPC (10 ms) according to the following simple expressions:

$$F_0(I_I{:}I_{II})_2 = F_0\left(I_I + \frac{1}{4}|I_I - I_{II}|\right),$$

$$F_0(I_I{:}I_{II})_3 = F_0\left(I_I + \frac{1}{2}|I_I - I_{II}|\right),$$

$$F_0(I_I{:}I_{II})_4 = F_0\left(I_I + \frac{3}{4}|I_I - I_{II}|\right), \quad \text{and analogously}$$

for $F_0(I_I{:}I_{IV})_{2,3,4}$ and $F_0(I_{II}{:}I_{IV})_{2,3,4}$.

Experiment 5

This experiment is an expansion of experiment 2, with the volume parameter from I transplanted to II, and then to IV. Each of the 5 different stimuli: I_S, II_S, IV_S, $II_{S(I)}$, $IV_{S(I)}$ was copied 5 times resulting in a total of 25 stimuli in the experiment.

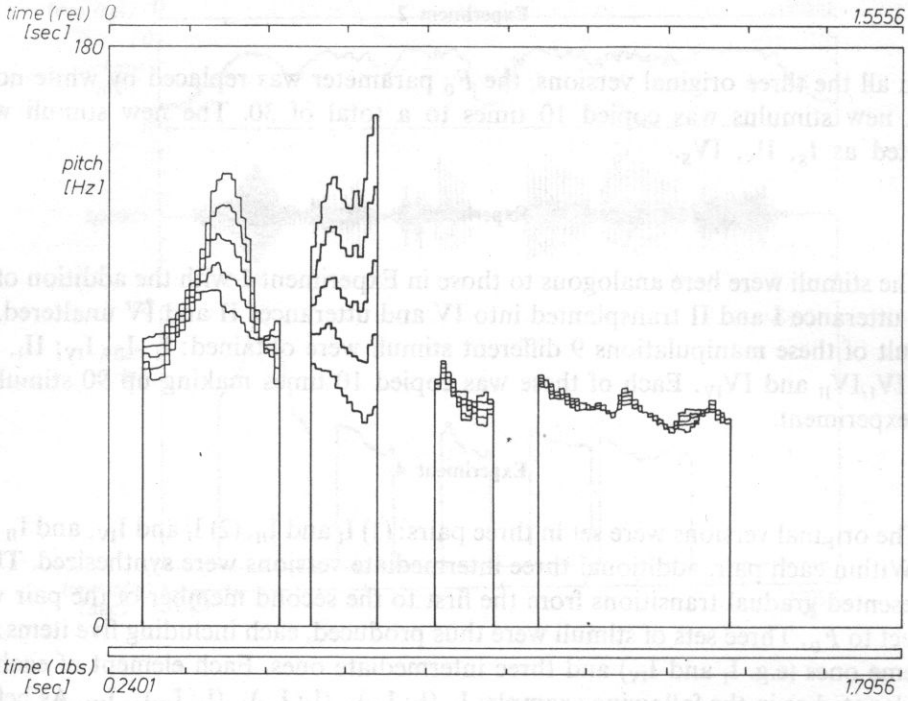


FIG. 2. F_0 in version I_I, I_{II} and three intermediate stimuli

Experiment 6

In all the three originals, the time-variable F_0 values were replaced by a constant $F_0 = 100$ Hz. 10 replications were produced.

Within each experiment, all final stimuli were randomized.

4. The listening test

The resynthesized experimental material described in the preceding Section was subjected to listening tests performed by three groups of subjects:

(1) Group (A) – 8 German students of phonetics at the University of Kiel. They had all had an experience of two or more years of phonetic training, but had no knowledge of the Polish language and were unfamiliar with the theory of Polish accentuation.

(2) Group (B) – 9 students of the A. Mickiewicz University of Poznań, all native Poles, specializing in various subjects, but with no knowledge of either phonetics or linguistics (apart from perhaps faint vestiges of such knowledge from primary or secondary schools).

(3) Group (C) – 10 research workers and technicians including subjects with and without previous phonetic training.

With one exception, none of the subjects in any of the three groups had any known hearing problem. The one exception belonged to Group (C). This subject had a unilateral hearing impairment, but his results did not evidently differ from those of the others in the Subgroup which was later distinguished on other evidence (the Systematic Group, see below).

The tests were carried out according to the same design in all three groups: Each subject received a set of answer sheets on which each stimulus, in each Experiment was marked by a successive number. Next to each number, there were five long horizontal dashes, each dash corresponding to one syllable, like this:

1. _____ _____ _____ _____ _____
2. _____ _____ _____ _____ _____
3. _____ _____ _____ _____ _____
⋮
40. _____ _____ _____ _____ _____
⋮

The listeners were asked to mark the syllable, or syllables, in each stimulus, which they heard as accented, according to the following istructions: If there is one accented syllable in a 5-syllable stimulus, mark it with a cross. If you hear two syllables as accented, you have a choice: If you hear both as being equally strongly accented, mark each of the two with a cross. If you hear one as being more strongly accented than the other, mark the former with a cross and the latter with a stroke. If you cannot make up your mind as to the location of the accent, make a random decision and mark any of the syllables (not more than just one) with a cross. The crosses were subsequently counted as primary accents, and the strokes as secondary accents.

5. Results of the tests

5.1. Homogeneity of responses

The arithmetic details of the analysis of response homogeneity are contained in Appendix A. We shall here confine ourselves to a general statement that Group (A) responded sufficiently systematically and consistently for us to regard it as

homogeneous. In Group (B), 3 subjects behaved consistently, and the others erratically in that the responses to identical stimuli within an Experiment were random. Consequently, only the former three subjects' responses were included in the subsequent statistical analysis. We would repeat here that (like the others in this Group) the three subjects had no previous phonetic training. Therefore, the drastic difference in the results should be attributed to such added psychological factors as motivation, concentration, interest, etc. In Group (C), again two subgroups could clearly be distinguished: 7 subjects behaved consistently and 3 erratically. This time, however, the first subgroup consisted of specialists – phoneticians, linguists and persons with experience with phonetic listening tests. The remaining 3 subjects did not have such qualifications.

Thus, from the original three groups, two final ones were created, *G* (German) and *P* (Polish) both homogeneous in the sense of consistency of responses. Group *P* included 10 subjests.

5.2. Statistical analysis of the consistent responses

The random variable is here the perceptual location of accent on the respective syllables. It comes in three values: primary accent denoted ***, secondary accent **, and no accent, denoted *. The distribution of the variable values has to be examined as an effect of three conditions assumed to be independent: (1) The subject's native language – two states: *G* and *P*; (2) the position of the syllable in the utterance – five states: syllable 1, ..., 5; (3) modification of the F_0 parameter in the re-synthesis: 3,4 or 5 states, depending on the kind of modification (or transplantation) of the F_0 parameter in the individual Experiments, as described in Section 3 above.

From the point of view of conditions (1) and (2), the results of the listening tests were analyzed separately for each state. Condition (3) was treated as the independent variable and the distribution of the random variable relative to the individual states of condition (3) was tested using a modified χ^2 test for independence in contingency tables. The design of the contingency tables may be illustrated by the following examples:

Experiment 1

Group *G*/syllable 1

	***	**	*
I	70	0	10
I_{II}	1	2	77
I_{IV}	6	13	61

Experiment 1

Group *G*/syllable 2

	***	**	*
I	3	6	71
I_{II}	79	0	1
I_{IV}	12	1	67

Experiment 1

Group *G*/syllable 3

	***	**	*
I	6	1	73
I_{II}	0	2	78
I_{IV}	3	3	74

Experiment 1

Group *G*/syllable 4

	***	**	*
I	3	6	69
I_{II}	4	11	65
I_{IV}	62	9	9

Experiment 1

Group *G*/syllable 5

	***	**	*
I	0	13	67
I_{II}	1	16	63
I_{IV}	0	3	77

As explained above, in the contingency tables, I_I refers to version I with unaltered parameters, I_{II} – to version I with F_0 transplanted from version II to version I and the remaining parameters unaltered, and I_{IV} – to version I with F_0 transplanted from version IV and the other parameters unaltered.

The results shown in the above contingency tables can be described as follows:

(1) Syllable 1 is strongly biassed towards *** (primary accent) under condition I_I and strongly biassed towards * (no accent) under conditions I_{II} and I_{IV}.

(2) Syllable 2 is strongly biassed towards *** under condition I_{II} and towards * under conditions I_I and I_{IV}.

(3) Syllable 3 is strongly biassed towards * under all three conditions.

(4) Syllable 4 is strongly biassed towards *** under condition I_{IV} and towards * under conditions I_I and I_{II}.

(5) Syllable 5 is strongly biassed towards * under all conditions.

It follows from the above observations that syllables 1, 2 and 4 are marked by the listeners as accented (primary accent) as an **effect of F_0** since the transplantation of F_0 from versions II and IV into version I, with the remaining parameters of version I unaltered resulted in the location of accent by the listeners as, respectively, in versions II and IV. Syllables 3 and 5 were perceived as unaccented no matter which of the three possible variations of F_0 the re-synthesized utterances contained.

Although the numbers in the contingency tables lead to an intuitively indubitable conclusion that the descriptions under points 1, ..., 5 above are correct, it may be interesting to define the statistical probability of the null hypothesis on the

non-dependence of the distribution of the three values of the variable under analysis on the conditions under (3) above (p. 9). Such non-dependence can be tested by using the χ^2 test. The application of this test to a full contingency table is, in its classical form, contingent on the individual cells including values not less than 8 (see, e.g., GREŃ [4]; 131) or – according to other authors – not less than 5 (see, e.g., VOLK [15]; 95). If this condition is not fulfilled, it becomes necessary, in the traditional usage, to combine classes in the contingency table, which leads to loss of information. As each of our tables above – like most the remaining ones – failed to meet the condition, a modification was introduced as described by NASS [11] and POTTHOFF & WITTINGHILL [12]. The appropriate algorithm and computer program for the modified χ^2 test was presented by DOBEK & KIEŁCZEWSKA [2], and the program was implemented on a computer of the JS system (Riad 32) in the Computer Centre of the Mickiewicz University in Poznań. The substance of the modified χ^2 test is presented here in Appendix B.

Experiments 1, ..., 5 were carried out with Group *G*, and Experiments 1, ..., 6 with Group *P*. For the former Group 45 contingency tables were constructed, and 48 for Group *P*. Each table was subjected to the modified χ^2 test. On the basis of the test values for each table, together with the corresponding value of the degrees of freedom (see Appendix B), the significance level for the rejection of the null hypothesis could be found in standard statistical tables.

In overall terms, for Group *G*, the null hypothesis of independence of the distribution of the random variable of the conditions (3) above, i.e., of the transformations of F_0, was rejected, for 2 tables, on a significance level of $\alpha = 0.01$, and for 19 tables it was impossible to give reliable significance level values because the modified number of degrees of freedom was less than 1. With respect to the remaining 24 tables, the null hypothesis was rejected at the level of $\alpha = 0.001$. In the case of Group *P*, the null hypothesis was rejected at $\alpha = 0.01$ once. Again, for 19 tables the number of degrees of freedom was less that 1, making it impossible to define the significance level, whilst for the remaining 28 tables the null hypothesis was, in this Group, rejected at $\alpha = 0.001$.

In both Groups, almost all the tables with the number of degrees of freedom below 1 pertained to syllables 3 and 5. Those are the **unstressed** syllables (i.e. the syllables without potential accent) /t͡ɕe/ and /ras/. In all three versions: I, II and IV, these syllables bore no accent. In these cases, the χ^2 test, or in fact any statistical test, is not really necessary, because the result is self-evident, as can be seen from the following typical examples:

Experiment 1

Group *P*/syllable 5

	***	**	*
I	0	1	99
I_{II}	0	0	100
I_{IV}	0	0	100

No comment is necessary on the dependence on condition (3) in a case like this. In other cases of syllable 3 or 5 the contingency tables were mostly as in the following example:

Experiment 2

Group *P*/syllable 5

	***	**	*
I	0	0	100
I_{II}	0	0	100
I_{IV}	0	0	100

5.3. Results of the listening tests

Experiment 1 (Fig. 3)

In Fig. 3, the horizontal axis refers to the successive syllables. The vertical axis indicates the percentage of the three possible reactions (i.e., ***, **, and *) out of a total of 80. The thick line connects the "primary accent" judgments, the thin line connects the secondary accent" judgments, while the dashed line pertains to "no accent". The three upper graphs refer to Group *G*. The successive graphs, both in the upper and the lower row, relate to I_I, I_{II} and I_{IV}. **The shift of the accent location** (primary accent) **from the first syllable** in I_I **through the second syllable** in I_{II} **to the fourth syllable** in I_{IV} is visible quite clearly.

Experiment 2 (Fig. 4)

Figure 4 as well as all the remaining Figures relating to the Experiments, are analogous to Fig. 3 with respect to the design and the meaning of the connecting lines. In Fig. 4, in both rows, the successive graphs refer to I_S, II_S and IV_S. The results differ as between the three levels of condition (3). There is, on the other hand, a fair similarity between the two Groups, *G* and *P*, though weaker than in Experiment 1. **The location of the accent, in this experiment, is evidently ambiguous.** For version I_S, there are two maxima at 1 and 4, for version II_S the maxima are at 2 and 4. Version IV_S exibits a strong maximum at 4 with a weaker one at 2. Thus, **the replacement of the variable F_0 by noise** (see p. 6) **considerably reduces the possibility of unambiguously locating accent.**

With respect to the secondary accent, the two Groups *G* and *P* differ

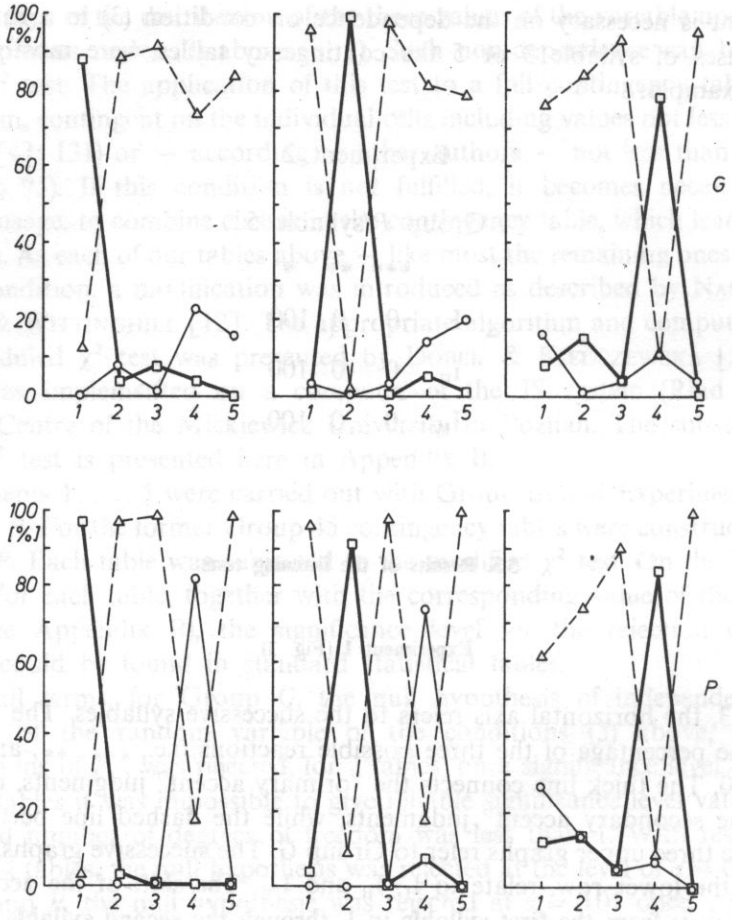


FIG. 3. Utterances I_I, I_{II} and I_{IV}. Results of the listening test in Experiment 1

considerably in Experiment 1 as well as in Experiment 2. In both Experiments the minima for primary accent (with the simultaneous maxima for no accent) lie at syllables 3 and 5.

Experiment 3 (Fig. 5)

Figure 5.1 refers to versions I_I, I_{II} and I_{IV}, Fig. 5.2 to versions II_I, II_{II} and II_{IV}, and Fig. 5.3. to versions IV_I, IV_{II} and IV_{IV}. All the three parts of Fig. 5 show very distinctly that the **location of accent is exclusively dependent on the course of F_0**, no matter whether the re-synthesized utterances were the originals (I_I, II_{II}, IV_{IV}) or had their F_0 transplanted from the other versions.

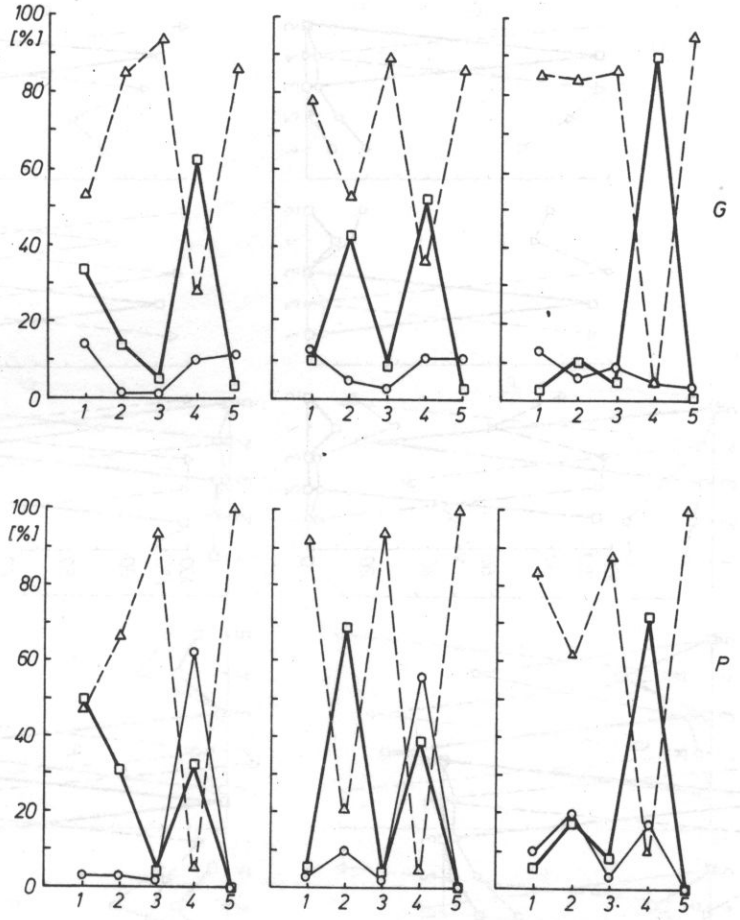
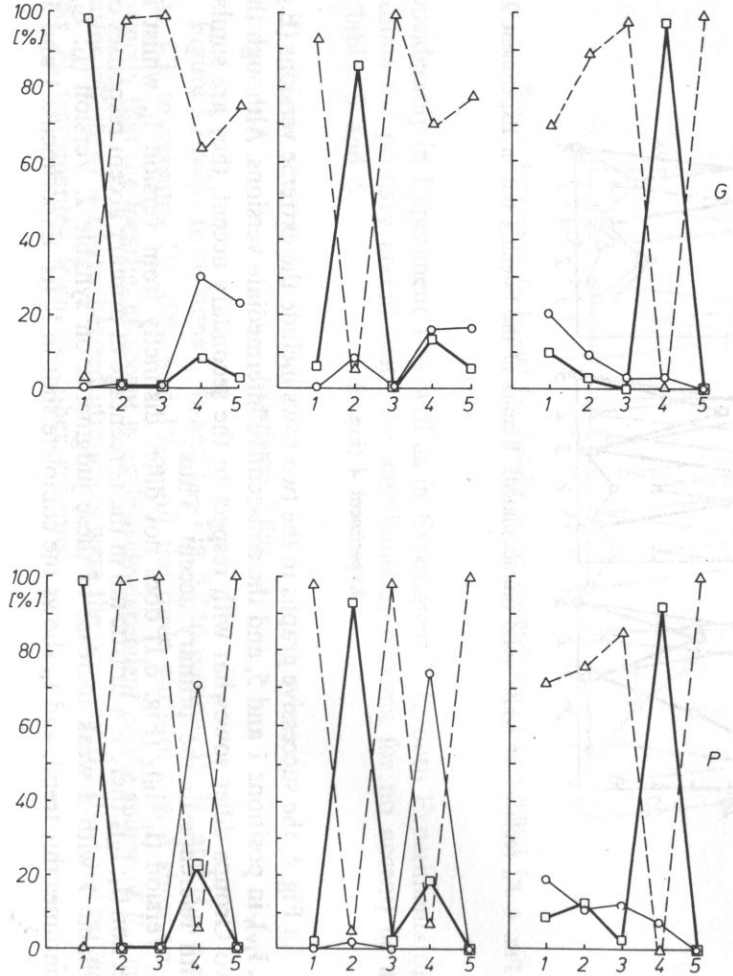


FIG. 4. F_0 replaced by white noise (unmodulated). Results of the listening test in Experiment 2

Experiment 4 (Fig. 6)

In Fig. 6, the successive graphs in the two rows include the extreme versions (E.g. I_I, I_{IV}), in positions 1 and 5, and the consecutive intermediate versions. Although the two Groups differ sowewhat with respect to the secondary accent, they are similar with reference to the primary accent. Thus:

Version $(I_I: I_{II})_2$ (Fig. 6.1) does not differ distinctly from version I_I, whilst in version $(I_I: I_{II})_3$ there is a drastic fall in the percentage of primary accent responses on syllable 1 with a weak increase of these judgments on syllable 2. Version $(I_I: I_{II})_4$ continues this trend, and I_{II} shows the complete shift in I_{II} (cf. Experiments 1 and 3).


100
[%]
80
60
40
20
0
1 2 3 4 5
G
P

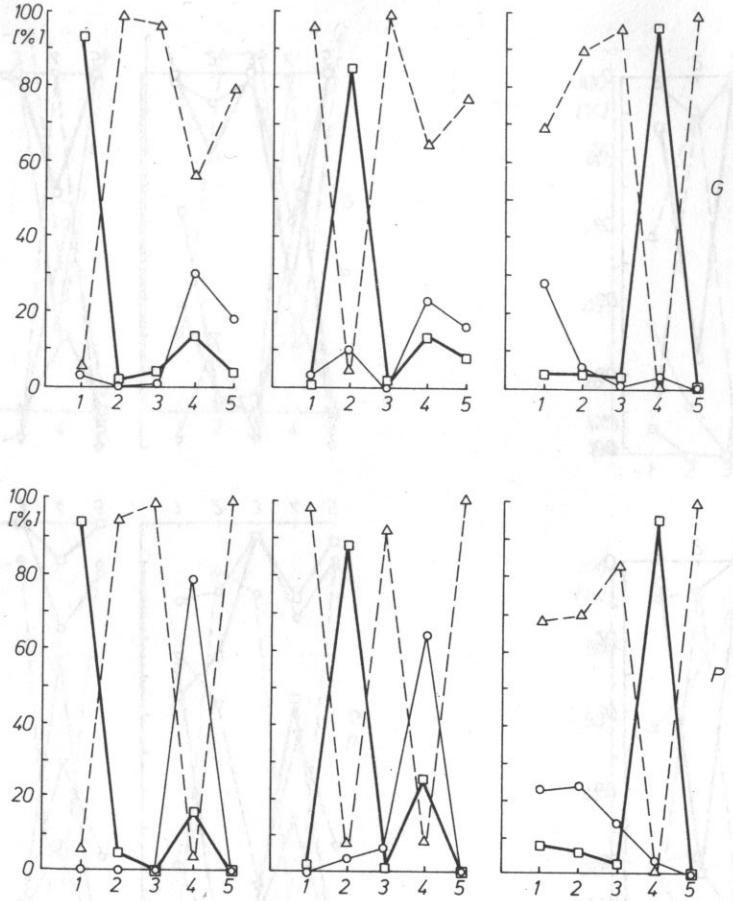
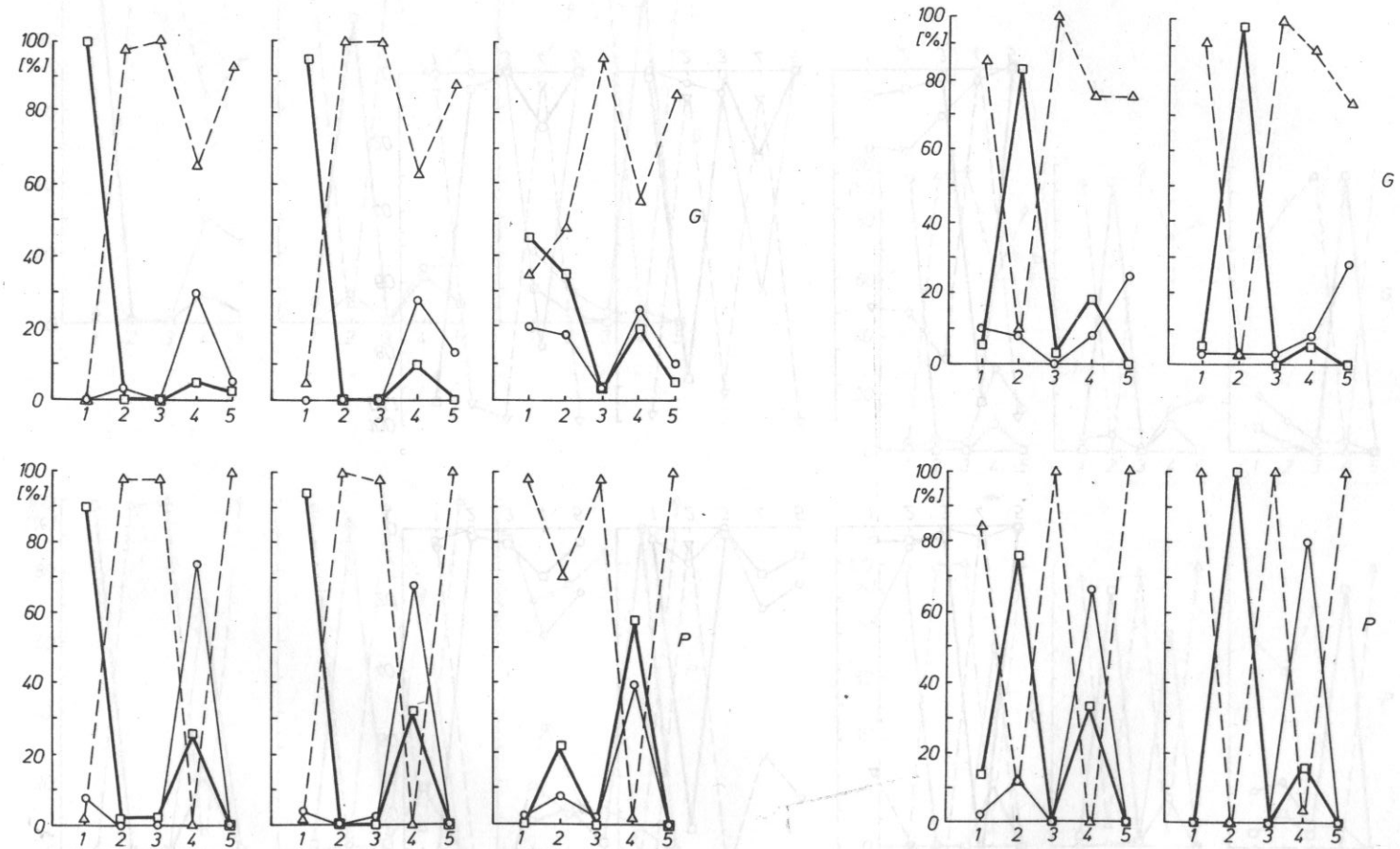
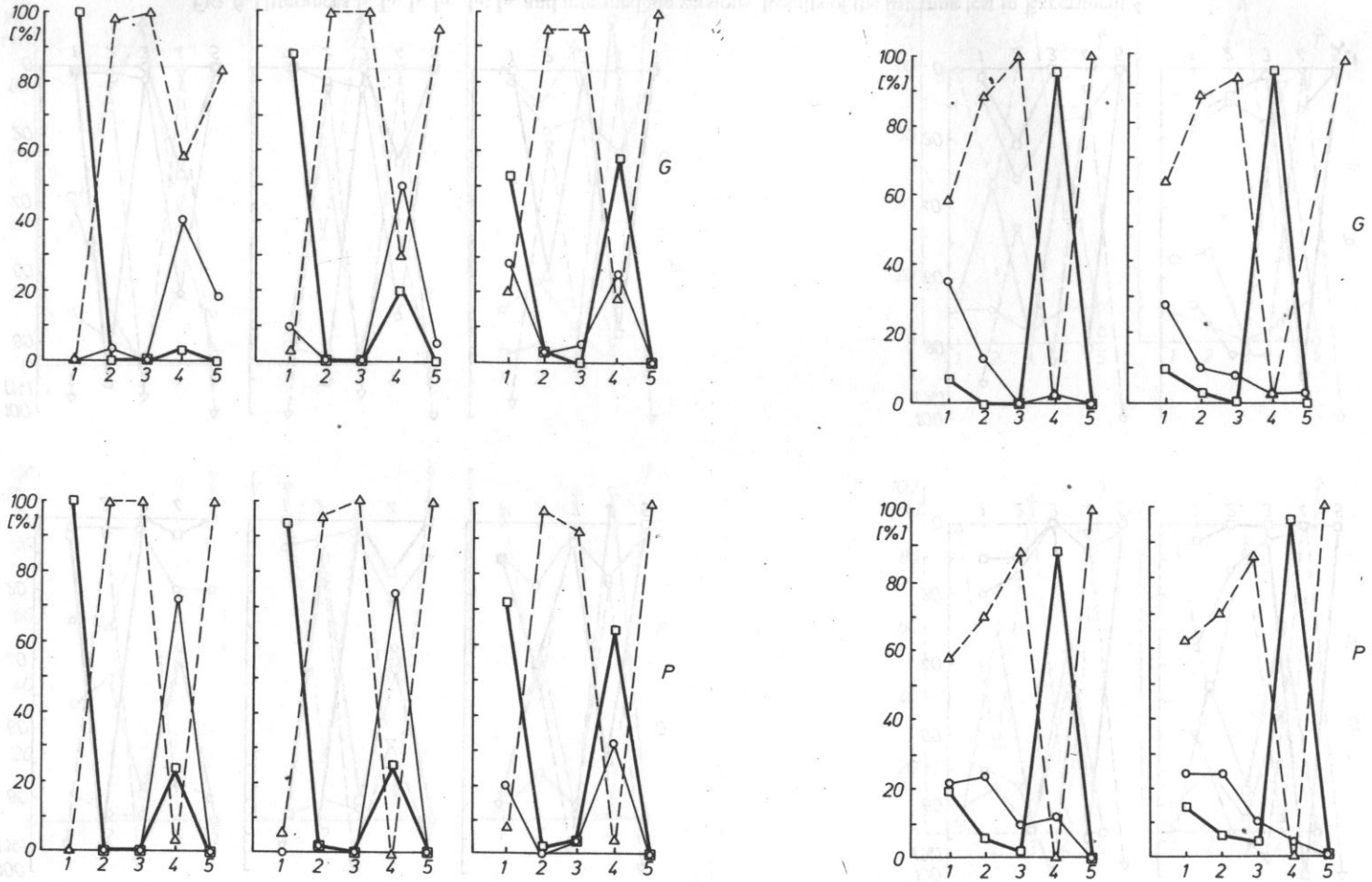


FIG. 5. Utterances I_I, I_{II}, I_{IV}; II_I, II_{II}, II_{IV}; IV_I, IV_{II} and IV_{IV}. Results of the listening test in Experiment 3


100
[%]
80
60
40
20
0
1 2 3 4 5
G
P


100
[%]
80
60
40
20
0
1 2 3 4 5
G
P

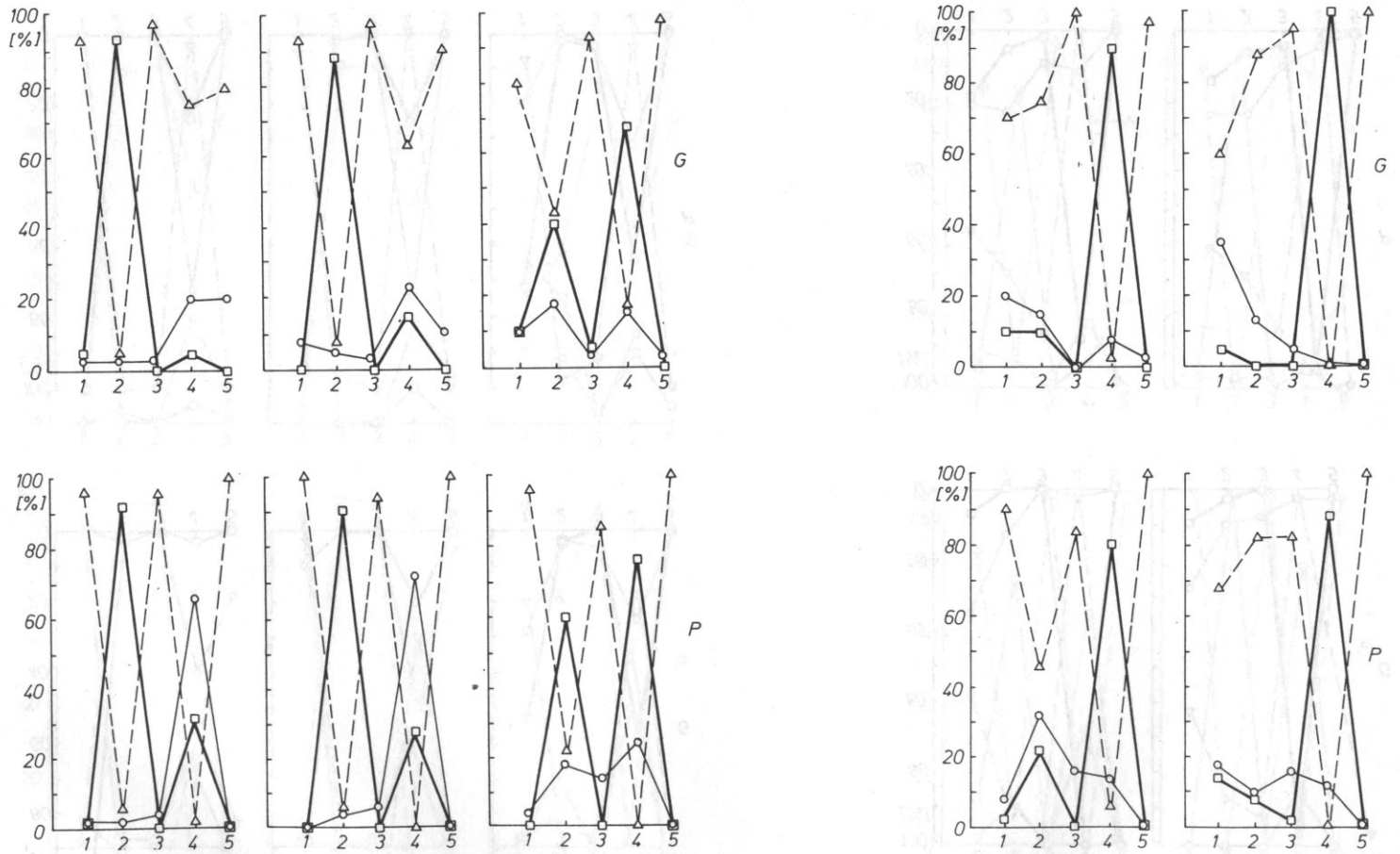


FIG. 6. Utterances I_I: I_{II}, I_I: I_{IV}, I_{II}: I_{IV} and intermediate versions. Results of the listening test in Experiment 4

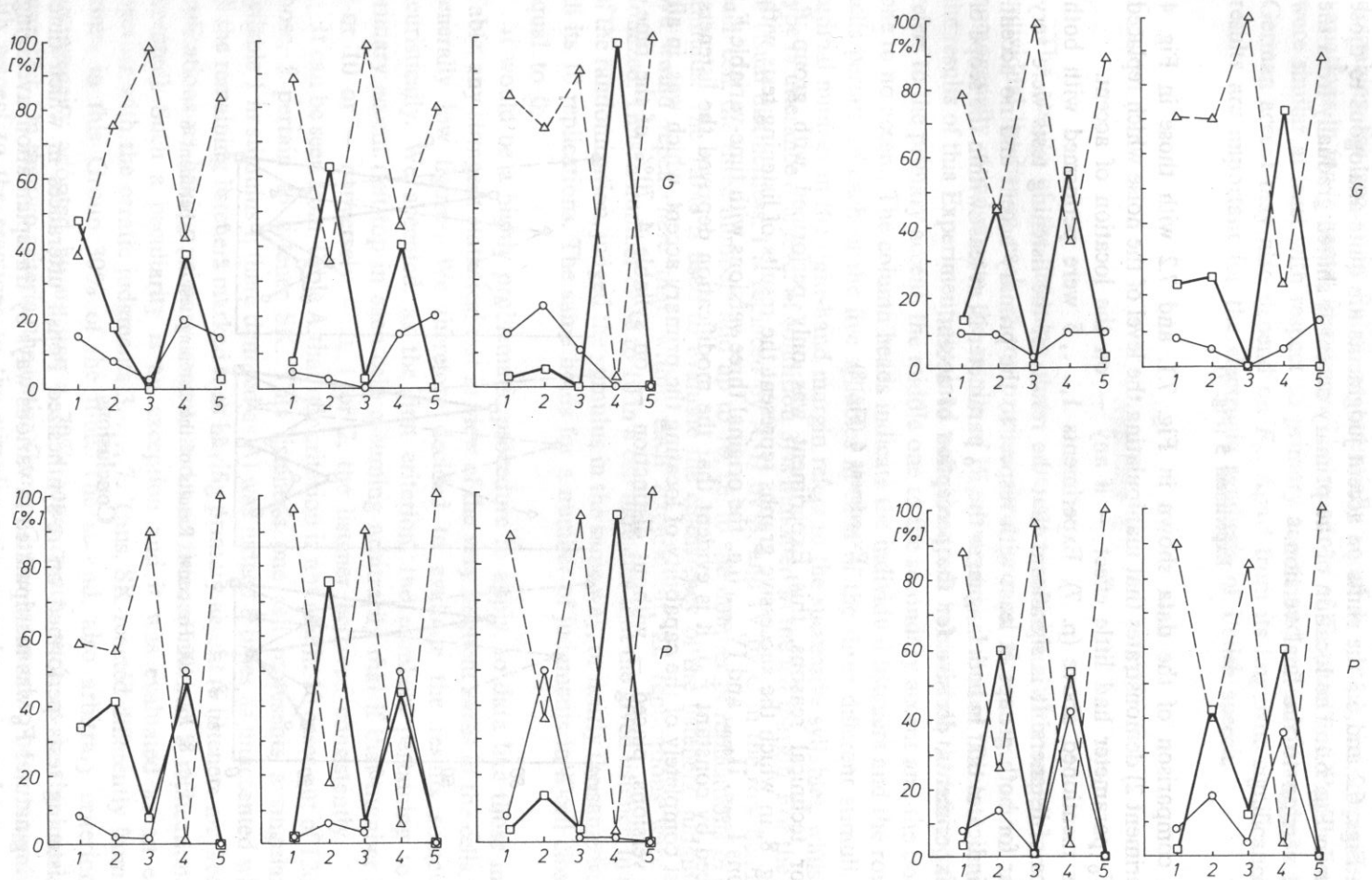


FIG. 7. F_0 replaced by noise (level-modulated). Results of the listening test in Experiment 5

In Figs. 6.2 and 6.3 the shifts on accent judgments are quite analogous to those shown in Fig. 6.1. The location of the primary accent is shifted gradually from the initial version to the final version.

Experiment 5 (Fig. 7)

A comparison of the data shown in Fig. 7.1. and 7.2 with those in Fig. 4 (Experiment 2) demonstrates that manipulating the level of the noise which replaced the F_0 parameter had little effect – if any – on the location of accent.

As mentioned above (p. 7), Experiments 1, ..., 5 were performed with both groups of listeners. It is significant that the results of the listening tests were very similar for both groups at least with respect to the primary accent and no accent. This indicates that **in both languages** the F_0 parameter is, at least within the scope of the Experiments, **decisive for the perception of accent.**

Experiment 6 (Fig. 8)

For technical reasons, this Experiment was only performed with group *P*. In Fig. 8, in which the successive graphs represent the results of listening tests with versions I_{const}, II_{const} and IV_{const} (i.e., the original three versions with time-variable F_0 replaced by constant F_0), it is evident that the modification deprived the listeners almost completely of the capability of locating the primary accent, which was, in all three versions, placed – without distinction – on syllable 4, 70% of the time.

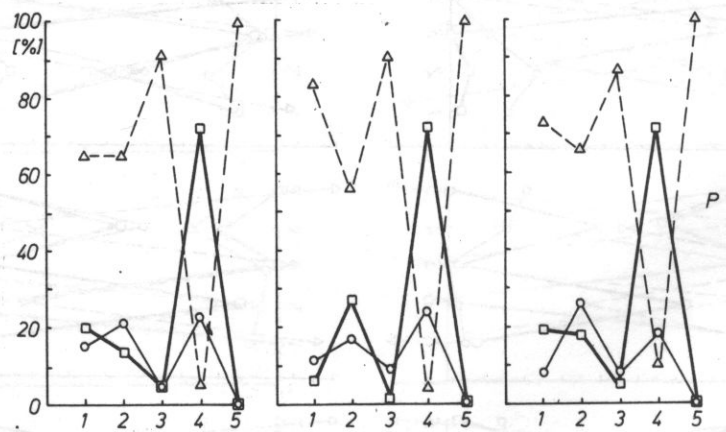


FIG. 8. F_0 = 100 Hz const. Results of the listening test in Experiment 6

Conclusions

Listening tests performed on re-synthesized Polish utterances in which time-varying natural F_0 was manipulated in various ways by transplantation, averaging, replacement by noise and converting to constant gave strong support to the theory

describing the accent in Polish as essentially tonal in nature, i.e., dependent on variations in fundamental frequency. The reactions of German and Polish listeners were similar at least with respect to primary accent and no accent suggesting that German accent may also depend on F_0. Apart from its linguistic significance, the results are important for the **electronic synthesis of Polish speech.**

Appendix A

The results of the listening tests were first examined with respect to each individual listener in each of the original three groups. For an evaluation of the homogeneity of judgments, the results of Experiment 1 were used. Table A sums up the results of this Experiment obtained with Group *G*. In each cell, the left-top figure refers to the primary accent, the middle one to the secondary accent and the bottom one to no accent. The column heads indicate the individual listeners and the rows of cells pertain to each of the five syllables in each of the three different stimuli. The ordinal number in the right-hand margin refers to the successive syllables, whilst the type of utterance (the level of condition 3) is indicated in the left margin. For the sake of visual distinctness, I is replaced by A, II by B, and IV by *C* in the row heads. In each cell, the three numbers add up to 10, i.e. to the number of replications of each stimulus. The Table is used to evaluate the **consistency** of the listeners' judgments. We would repeat here that, within each Experiment, all replications were randomized. Thus, a number close to 10 in a cell indicates that the given listener, in spite of the randomization, judged the stimulus in the same way, or nearly the same way in all its 10 replications. The same holds for a number of judgments in a cell close or equal to 0.

It would be a highly problematic procedure to apply, to data like those in the Table, any stringent statistical test in view of the very frequent zeros in the cells, and generally low figures. We therefore decided to evaluate the results somewhat heuristically. We accepted, as the first criterion, the numbers pertaining to the primary accent (left-top in each cell) assuming arbitrarily that if that number is 8, 9 or 10 or — conversely — 0, 1 or 2, the listener behaved consistently.

It can be seen from Table A that our criterion is not met in 10 cases out of 120. Of those, 5 pertain to listener SK. This listener's one cell represents a singularity: Syllable 1 in stimulus I_I (here denoted as A) was judged 8 times as unaccented whilst all the remaining listeners marked it as having primary accent (6 listeners marked all replications as having primary accent, and one listener marked 8 replications as accented). Such a peculiarity is an exception and it was evaluated as an error together with the erratic judgments 3, ..., 7. Thus, SK reacted differently from the others in this Group 30% of the time. The second, also arbitrary criterion of homogeneity adopted by us was a 70% concurrence, such as in the case just described. On the two criteria, the entire Group *G* was assumed to be homogeneous with respect to the reactions to the stimuli.

	BK	CR	EB	FR	HW	MT	RD	SK	
A 2	10	10	10	8	10	10	10	2	
1	0	0	0	0	0	0	0	0	
0	0	0	0	2	0	0	0	8	
A_B 2	1	0	0	0	0	0	0	0	
1	1	0	0	0	0	0	1	0	1
0	8	10	10	10	10	10	9	10	
A_C 2	0	0	0	0	1	0	5	0	
1	2	0	3	3	1	1	3	0	
0	8	10	7	7	8	9	2	10	
A 2	0	0	0	0	0	0	0	3	
1	1	0	0	0	3	0	2	0	
0	9	10	10	10	7	10	8	7	
A_B 2	10	10	10	10	10	10		9	
1	0	0	0	0	0	0	0	0	2
0	0	0	0	0	0	0		1	
A_C 2	6	0	2	0	2	0	0	2	
1	0	0	0	0	0	0	0	1	
0	4	10	8	10	8	10	10	7	
A 2	0	0	0	2	0	0	0	4	
1	0	0	0	0	1	0	0	0	
0	10	10	10	8	9	10	10	6	
A_B 2	0	0	0	0	0	0	0	1	
1	1	0	0	0	0	0	0	1	3
0	9	10	10	10	10	10	10	8	
A_C 2	2	0	2	0	2	0	0	3	
1	0	0	0	0	0	0	0	0	
0	8	10	8	10	8	10	10	7	
A 2	0	0	0	1	0	0	0	2	
1	0	0	5	1	0	6	1	3	
0	10	10	5	8	10	4	9	5	
A_B 2	1	0	0	0	0	2	0	2	
1	0	0	7	0	0	2	1	1	4
0	9	10	3	10	10	6	9	7	
A_C 2	4	10	8	10	7	10	6	7	
1	3	0	2	0	0	0	4	0	
0	3	0	0	0	3	0	0	3	
A 2	0	0	0	0	0	0	0	0	
1	4	0	0	0	7	0	2	0	
0	6	10	10	10	3	10	8	10	
A_B 2	1	0	0	0	0	0	0	0	
1	1	0	2	0	5	0	8	0	5
0	8	10	8	10	5	10	2	10	
A_C 2	0	0	0	0	1	0	0	0	
1	1	0	1	0	1	0	0	0	
0	9	10	9	10	8	10	10	10	

Table A. Reactions of the German listeners in Experiment 1

The same, informal test was applied to Group *B* and it was found that on the above combined criterion only 3 subjects behaved consistently, so the remaining 6 were eliminated from further analysis as being outside the homogeneous panel.

In Group *C* consisting of 10 subjects, 3 were eliminated on the same basis. The three subjects from Group *B* were then combined with the three subjects of Group *C* obtaining a homogeneous group of consistent respondents, and this new Group was denoted by *P*. As in the case of Group *G* (the entire panel of German students), the homogeneity test was based on Experiment 1.

As none of the Polish students (the original Group *B*) had any phonetic training, whilst – on the other hand – some of the subjects belonging to Group *C* had at least had some experience with phonetic experiments, our results permit no generalization as to the effect of such experience on the results of phonetic listening tests. But

a comparison of the results of all the three original Groups strongly suggests that such factors as motivation, concentration and other psychological conditions, which have so far received little attention in phonetic literature, may be significant.

Appendix B

The procedure applied by us here in order to analyze statistically the contingency tables does not put any constraints on the numbers in the individual cells apart from an obvious case like the one quoted above with reference to Experiment 2/Group *P*/syllable 5 (p. 13).

Symbolically, a contingency table may be represented as follows:

$$
\begin{array}{llll}
x_{11} \ldots\ldots\ldots\ldots x_{1j} \ldots\ldots\ldots\ldots x_{1m} \ldots\ldots z_1 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_{i1} \ldots\ldots\ldots\ldots x_{ij} \ldots\ldots\ldots\ldots x_{im} \ldots\ldots z_i \\
x_{n1} \ldots\ldots\ldots\ldots x_{nj} \ldots\ldots\ldots\ldots x_{nm} \ldots\ldots z_n \\
\hline
y_1 \qquad\qquad y_j \qquad\qquad y_m \qquad\qquad N
\end{array}
$$

In the above table, x_{ij} denotes the number of observations in the i-th row and j-th column, i.e., the size of the sub-class (i, j) in a double classification with the value of z_i as a marginal sum for the i-th row and y_j as the marginal sum for the j-th column. The probability of the sub-class (i, j) is defined by the corresponding marginal sums and the grand sum N.

The null hypothesis on the independence of the features may be written as follows:

$$H_0: = p_i p_j \quad \text{for} \quad i = 1, 2, \ldots, n \quad \text{and} \quad j = 1, 2, \ldots, m.$$

The classical test of this hypothesis is based on the statistic

$$G = N \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij}^2/(y_j z_i) - 1,$$

where

$$N = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij},\ y_j = \sum_{i=1}^{n} x_{ij},\ z_i = \sum_{j=1}^{m} x_{ij}.$$

If the null hypothesis is true, then the statistic G has a distribution which is approximately that of χ^2 with $\nu = (n-1)(m-1)$ degrees of freedom.

If the numbers in the individual cells are small, failing to meet the classical condition (see above p. 12), the test is subject to modification. The G statistic is

replaced by the G_1 statistics:

$$G_1 = cG, \quad \text{where}$$

$$c = 2E(G)/\mathrm{Var}(G),$$

where $E(G)$ is the expected value of G and $\mathrm{Var}(G)$ is the variance of the statistic G. The calculation of the value of the G_1 statistic is sowhat complex. The appropriate numerical methods are presented in DOBEK & KIEŁCZEWSKA ([2]). The procedure calculates the modified χ^2 value as well as the modified value of the number of degrees of freedom:

$$\nu_1 = c(m-1)(n-1)/(N-1).$$

References

[1] W. J. BARRY, W. van DOMMELEN, H. JANSSEN, K. J. KOHLER, K. SCHÄFER W. THON, G. TIMMERMAN, *Phonetic data processing*, Kiel University, Arbeitsberichte Nr. 18, Inst. f. Phonetik 1982.

[2] A. DOBEK, H. KEIŁCZEWSKA, *Testing independence and homogeneity of samples for small expected frequencies* (in Polish). Algorytmy biometryczne i statystyczne, 6, 163–174 1979.

[3] K. DOBROGOWSKA, *Investigation of the relations between accent and intonation in unemotional speech* (in Polish). Lingua Posnaniensis **21**, 65–75 (1978).

[4] J. GREŃ, *Mathematical statistics, models and exercises* 4th *ed* (in Polish) PWN, Warszawa 1974.

[5] W. JASSEM, *The phonology of Polish stress*, **15**, 252–269 (1964).

[6] W. JASSEM, *The accent in Polish* (in Polish) Prace Językoznawcze No. 31, Komitet językoznawstwa PAN, 1962.

[7] W. JASSEM, D. GIBBON, *Re-defining English accent and stress*, Journal of the International Phonetic Association **10**, 2–16 (1980).

[8] W. JASSEM, J. MORTON, M. STEFFEN-BATOGOWA, *The perception of stress in speech-like stimuli by Polish listeners*, in: Speech analysis and synthesis ed. W. Jassem PWN, Warszawa 1968 vol. 1 pp. 289–308.

[9] P. LADEFOGED, *A course in phonetics*, Harcourt, Brace, Jovanovich Inc., New York 1975.

[10] O. LANCASTER, *The chi-squared distribution*, Wiley and Sons, New York 1969.

[11] C. A. D. NASS, *The χ^2 Test for small expectations in contingency tables with special reference to accidence and absenteeism*, Biometrika **46**, 365–385 (1959).

[12] F. POTTHOFF, M. WITTINGHALL, *Testing for homogeneity. The binomial and multinomial*, Biometrika **53**, 167–182 (1966).

[13] G. RATHJEN, *Numerische und experimentalle Methoden der lineären Prädiktion zur Datenreduktion von Sprachsignalen* Arbeitsberichte Nr 11, Inst. f. Phonetik, Kiel 1979, 3–133.

[14] K. SCHÄFER, *Concepts in the SSP Programme*, Arbeitsberichte Nr. 18, Inst. f. Phonetik, Kiel 1982, 11–126.

[15] W. VOLK, *Mathematical statistics for engineers*, (in Polish) Wyd. Nauk.-Techn, Warszawa 1965.

[16] B. WIERZCHOWSKA, *Polish pronunciation* (in Polish) Państw. Zakł. Wyd. Szkol. Warszawa 1965.

[17] B. WIERZCHOWSKA, *A phonetic description of Polish* (in Polish) PWN, Warszawa 1987.

[18] B. WIERZCHOWSKA, *The phonetics and phonology of Polish* (in Polish), Zakł. Nar. im. Ossolińskich, Wrocław 1980.

Received on September 12, 1989